Low Power Constrains the Space of Narratives Available to Speakers

Judy Sein Kim (judyseinkim@gmail.com)
Department of Psychology & University Center for Human Values
Princeton University

M.J. Crockett (mj.crockett@princeton.edu)
Department of Psychology & University Center for Human Values
Princeton University

Abstract

How does power shape communicative freedom? During communication, speakers often balance multiple competing goals—such as being informative versus preserving their reputation. Across four experiments (N=~1400), we examined how the power relation between speakers and listeners affects narrative production in the face of conflicting informational and reputational goals. Participants imagined having to confess about a minor wrongdoing and were assigned to low, equal, or high power roles (an employee speaking to a supervisor, co-worker, or intern). Low power speakers were more likely to prioritize informativeness and believability. Importantly, the open-ended narratives they provided were more homogenous than those of high or equal power narrators, suggesting that power restricts the range of utterance choices considered. Speakers' ratings of pre-written narratives indicated that power may additionally affect how speakers evaluate the believability of different utterance choices.

Keywords: Language, Production, Power, Communication RSA, Narratives

Introduction

How do we decide what to say? Human communication is flexible: there is no rule book that tells us exactly what words to use in a given situation. Rather, people are typically free to make different decisions about what to say, even in the same context.

Variation in utterances can come from speakers adopting different strategies to balance multiple, often conflicting, goals (e.g., Goodman & Frank, 2016; Yoon et al., 2020; Degen, 2023). One common situation is when a speakers' duty to be informative (truthful, relevant, concise, etc.) conflicts with a desire to preserve their reputation (e.g., Kim, Colombatto & Crockett, 2024; Kim & Crockett, under review; Banerjee, Heyman & Lee, 2020; Sedikides, Hoorens & Dufner, 2015). For example, when telling the truth means revealing undesirable information about oneself (e.g., admitting wrongdoing), speakers have to decide how to manage conflicting *informational* and *reputational goals*. They might tell the truth and sacrifice their reputational goals, or lie and compromise their informational goals. A speaker might even try to balance both goals by doing something in-between telling the truth and lying (e.g., by revealing a partial truth, evading, implying a lie, being indirect; Clementson, 2018; Camp, 2018; Rogers et al., 2017; Weissman & Terkourafi, 2019; Yoon et al., 2020).

The Rational Speech Act model (RSA) provides a useful framework for further articulating the relationship between speakers' goals and their utterance choices (Goodman & Frank, 2016). According to this framework, speakers consider a set of alternative utterances, evaluating how each choice achieves the relevant goals (e.g., determining their informational and reputational values) and deciding which goals to prioritize. Speakers choose the utterance highest in utility (e.g., a weighted sum of informational and reputational values).

Following this framework, we recently showed that speakers' choices to tell truthful, false, or under-informative narratives about their moral actions can be explained through different strategies for managing conflicting informational and reputational goals (Kim & Crockett, under review). In the current paper, we examine the effect of *power differentials* between speakers and listeners on various components of narrative production, including weights placed on goals, valuation of utterances, and the range of alternative utterances considered. We focus on the effect of power on narrative production in a particular setting: a workplace hierarchy.

A basic definition of power is the ability to act, or the 'power-to' do things (e.g., Hobbes, 1985; Pitkin 1972, Morriss, 2002). Some people have more 'power-to' than others: in the workplace, for example, a regular employee has a more restricted range of actions available to them than a CEO. In terms of communicative power-to, a well-documented phenomenon is that of self-silencing, where low-power employees do not freely offer their ideas and opinions, especially if they are challenging or critical (e.g., Tangirala & Ramanujam, 2008a; Van Dyne et al., 2003; Morrison, 2023; Pfrombeck et al., 2022; Mazzei et al., 2025). Employees might not exercise their voice due to various reasons, such as fear that volunteering unsolicited opinions crosses some boundary given their position or a belief that doing so will be futile (e.g., Janssen & Gao, 2015; Frazier & Bowler, 2015). Put differently, one effect of low power is that it prevents speakers from volunteering information that they otherwise could with more power (i.e., cause them to become non-informative).

In some situations, speakers may not want to speak truthfully, but do not have the option of withholding information. For example, suppose a manager asks an employee for their opinion on a project, but the employee's opinion is negative and could potentially offend the manager. Given the direct inquiry, silence is less available as an option. Any speaker in this situation (regardless of how much power they have) then has at least three options: tell the undesirable truth, lie, or do something in between like indirectly voicing

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*

their opinion, allowing them to avoid stating the truth explicitly but not lying (e.g., Brown & Levinson, 1987; Lakoff, 1989; Harris, 2003; Yoon et al., 2020). Speakers who choose to explicitly lie or give an indirect answer de-prioritize informational goals in different ways: lying means the speaker does not provide truthful information, while indirectness means not giving clear, relevant information.

Importantly, some studies of communicative exchanges between speakers and listeners with power differentials suggest that the effect of low power is to push speakers to de-prioritize their informational goals, in both ways. Subordinates in real workplaces report deceiving their superiors more than vice versa (e.g., Lindsey et al., 2011; Deluga, 1995). People with low power also tend to be more polite, including by being indirect, to those with higher power than vice versa (e.g., Morand, 2000, 2014; Brown & Levinson, 1987; Holtgraves, 1997).

Another important aspect of power in this context is that of having 'power-over' another person, often through control over resources, punishment, and rewards (e.g., Weber, 1978; Dahl, 1957; Fiske, 1993; Keltner, Gruenfeld & Anderson, 2003; Thibaut & Kelley, 1959). This type of power is relational and inherent to workplace hierarchies: with low power-over, an employee might fear punishment from the manager for honestly providing unpleasant information, but the opposite is much less likely. In this way, low power-over, or increased fear of consequences for telling the truth, explains why low power speakers are more likely to lie, be under-informative, or not speak at all (i.e., place low weight on informational goals).

In this paper, we test three other predictions about what low power speakers compared to high power speakers might do in the face of conflicting informational and reputational goals: they may 1) be more limited in the alternatives they consider, 2) place higher on weight informational goals, as well as on presentational goals of being believable, and 3) evaluate narratives as having less presentational value.

First, we propose that the constraining effect of low power should be observable in the range of narrative options that speakers consider at all. This follows from the observation that for low power speakers, certain narrative options become riskier and less ideal than if they had more power. We test this prediction by comparing the variability of narratives produced by speakers with more or less power, with the assumption that the range of narratives produced by speakers within a group can be a proxy for the alternatives available to an average speaker in the group.

Second, we demonstrate that low power does not always lead speakers to compromise on their informational goals. In some contexts, low power may compel speakers to be *more* informative than they would be if they had more power (resulting in what McKinney, 2016 calls 'extracted speech'). One situation where this might occur is if there are consequences for being found out after not providing an informative answer. In such a situation, low power speakers should be particularly worried about being found out as deceptive, as the consequences may be harsher. We therefore predict that if the risk of being suspected or found out as a liar is a salient risk, low power speakers should be *more* likely to be fully informative.

In fact, given the possibility of being found out, we propose that an additional, second-order *presentational goal* may become particularly important, where speakers should strive to make the listener believe that they are being as informative as possible, and avoid suspicion of dishonesty (Kim & Crockett, 2022; Kim & Crockett, under review; see also Yoon et al., 2020). While all speakers should want to be seen as honest (i.e., having presentational goals), low power might heighten a speaker's presentational concerns. This predicts that low power speakers may additionally place high weights on presentational goals.

Finally, we test the possibility that power affects the valuation of narratives in terms of how well they meet a speaker's goals. Power can be instantiated in structural forces, like in social norms of discourse (e.g., Foucault, 1971; Bourdieu, 1977). Norms about power and language may make some narratives (e.g., under-informative narratives) perceived as more or less acceptable and normal depending on the speaker's power (e.g., Holmes & Stubbe, 2015). Speakers reasoning about such conversational norms might therefore evaluate the value of narratives differently, particularly for presentational value. Given the same narrative, a low power speaker may be more likely to think that a norm-breaking narrative will arouse suspicion. Relatedly, speakers with low power may be given less credibility from listeners in a prejudicial way, where they are less believed than if they had more power (Carney et al., 2013; Carmona, 2023; Dotson, 2011; Fricker, 2007; Medina, 2011; Davis, 2016; Ulatowska & Cislak, 2022). Speakers with low power who infer that they are less likely to be believed, would also assign lower presentational value to the same narrative choices than their high power counterparts.

In a set of experiments, participants (henceforth, "speakers") read a scenario about being an employee at a company who steals snacks from the break room one night. Speakers were asked to imagine being in the situation described, and imagine having to respond to another person's inquiry about their behavior, setting up a conflict between informational and reputational goals, as well as a possibility of being found out. Speakers were assigned to one of three power groups: low power (where the listener is their supervisor), equal power (listener is a co-worker), and high power (listener is their unpaid intern). The equal power condition was included to examine whether any effects of power are due to having more power (predicting that high power narratives would stand out), or less power (predicting that low power narratives would stand out).

To test the hypothesis that power affects the range of alternative narratives available, we first elicit open-ended narratives, comparing the diversity of responses across groups of speakers with different power (Experiments 1a-d). We test the prediction that in managing conflicting informational and reputational goals, speakers with lower power will produce more homogenous narratives than those

with high power. To explore how and why particular kinds of narratives might be favored, in some experiments (Experiments 1b-d, 2a) speakers were additionally presented with pre-written choices. For both kinds of narratives (written by the participant vs. presented in the experiment), speakers rated their informational, reputational, and presentational values. We predicted that in general, speakers with low power relative to those with high power would *weigh* their informational and presentational goals more in their narrative choices, and avoid narratives with lower presentational values. We further asked whether low power speakers compared to their high power counterparts assign lower presentational values to the same pre-written narrative choices.

Open-ended Narratives (Experiments 1a-d)

Participants

All participants were recruited on Prolific (www.prolific.com). Within each experiment (1a-d), participants were randomly assigned into a power group (Exp 1a: Low n=50, Equal n=50, High n=48; Exp 1b: Low n=100, Equal n=99, High n=99; Exp 1c: Low n=147, Equal n=151, High n=149; Exp 1d: Low n=206, Equal n=199, High n=196). Participants were fluent English speakers from the US, gave consent according to Yale University and Princeton University IRBs, and paid $7.5-$12 per hour.

Experimental Stimuli & Procedures

Participants ("speakers") first read a vignette where they were asked to imagine being an employee who steals snacks from their office break room, (refer to Kim & Crockett, under review, for the full text). In the vignette, they come back to their office at night to grab their phone charger, see shared snacks lying around in the break room, and decide to take them all home. The next day, during a meeting, someone ("listener") turns to them and says that there was an email saying all the snacks are gone. They ask, "Weren't you just telling me that you were here last night?" Speakers were then asked to imagine responding to the listener and write their response.

For the power manipulation, vignettes only differed in who the listener was (Low: the speaker's supervisor, Equal: speaker's co-worker, High: speaker's unpaid intern). Speakers were then asked to rate their own narrative in terms of their informational ("How well does your response reflect the truth?"), reputational ("Will the [supervisor/co-worker/intern] think that you're a bad person?"), and presentational values ("Will the [supervisor/co-worker/intern] suspect that you're hiding something or not telling the full truth?"). All questions were rated on a scale of 'Not at all' to 'Very much'.

Different versions of the experiment (Exp1a-d) had slightly different wordings for questions or contained additional questions that are not analyzed for this paper.

Results

Low power narratives are more homogenous

Our main prediction was that power would have an effect on the range of narratives that are considered. To investigate this, we looked at how similar narratives were to each other within each power group. We measured within-group variance in narratives using bag-of-words (BOW) cosine similarity. Narratives were first minimally cleaned for spelling and typos and words were lemmatized.

As predicted, there was a significant effect of power on within-group cosine similarity, and this replicated across all experiments (**Fig. 1**; linear mixed effects model with low power as the reference level). In all experiments, low power narratives were more similar to each other than high power narratives (E1a: B=-0.06, SE=0.01, F(2,297)=4.17, p<.001; E1b: B=-0.05, SE=0.01, F(2,296)=4.94 p<.001; E1c: B=-0.05, SE=0.01, F(2,444)=5.18, p<.001; E1d: B=-0.05, SE=0.01, F(2,598)=5.48, p<.001). In all but one experiment (1d), low power narratives were also more homogenous than equal power ones (E1a: B=-0.05, SE=0.01, F(2,297)=4.07, p<.001; E1b: B=-0.04, SE=0.01, F(2,296)=3.3, p=.001; E1c: B=-0.04, SE=0.01, F(2,444)=4.59, p<.001; E1d: B=-0.01, SE=0.01, F(2,598)=1.13, p=.257).

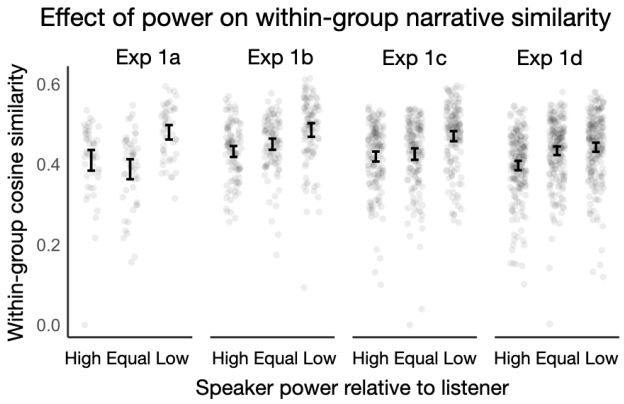

Figure 1. The effect of power on narrative similarity (within-group cosine similarity). Error bars are mean +/- 95% confidence intervals.

Low Power narratives have higher informational and presentational values

Our second prediction was that low power speakers would be more concerned about informational and presentational values of narratives, at the cost of lower reputational values. We next examine speakers' ratings of their own narratives in terms of informational, reputational, and presentational values.

Since the previous result suggested that high and equal power narratives are similar, for simplicity we only compare high vs. low power narratives from here on. In two out of four experiments, low power narratives were rated more highly in terms of their informational and presentational values

(**Figure 2**, top; LMEM within each experiment and question, effect of power on ratings, for informational value: E1a: B=0.35, SE=4.4, F(1,199)=0.006, p=.94; E1b: B=-4.09, SE=5.21, F(1,198)=0.62, p=.43; E1c: B=-12.51, SE=4.59, F(1,294)=7.42, p=.0068; E1d: B=-9.1, SE=3.34, F(1,400)=7.21, p=.0076; for presentational value: E1a: B=-9.1, SE=3.34, F(1,199)=0.4, p=.53; E1b: B=-1.68, SE=3.53, F(1,198)=0.23, p=.635; E1c: B=-9.06, SE=3.17, F(1,294)=8.18, p=.0045; E1d: B=-6.84, SE=2.94, F(1,400)=5.42, p=.02). However, power did not have an effect on reputational value ratings.

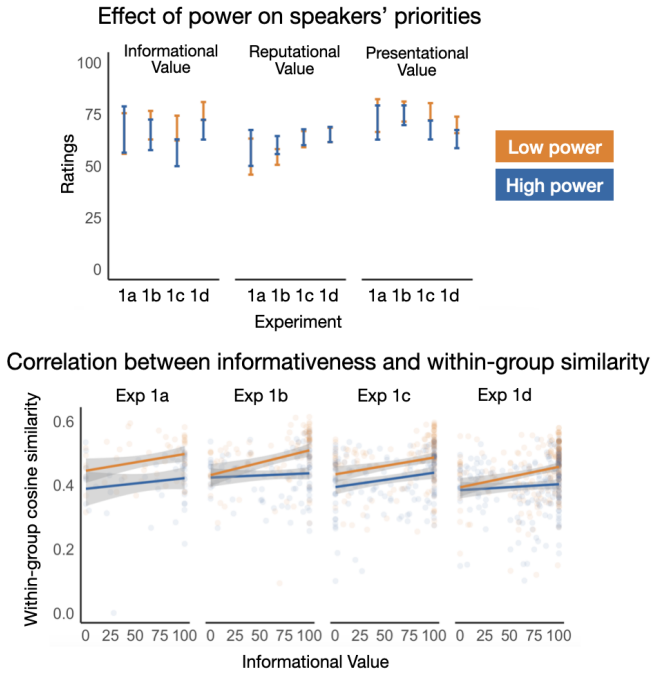

Figure 2. Top: Effect of power on the informational, reputational, and presentational values of speakers' own narratives. Bottom: Effect of power on correlation between informational value and within-group narrative similarity.

Since our prediction about narrative similarity is that low power speakers will be more restricted to being informative, we correlated each narratives' informational value with its within-group cosine similarity score (**Fig. 2**, bottom). Consistent with our predictions, in the low power group, informational value was always positively correlated with within-group narrative similarity, suggesting that low power narratives may indeed be more similar to each other because speakers are trying to be informative (E1a: r^2=0.08, p=.04; E1b: r^2=0.11, p<.001; E1c: r^2=0.07, p=.002; E1d: r^2=0.06, p<.001). In other words, in the low power group, less informative narratives are ones that deviate from other narratives in the group. In contrast, for the high power group, there was no relationship in three out of four experiments (E1a: r^2=0.002, p=.3; E1b: r^2=0.005, p=.5; E1c: r^2=0.05, p=.008; E1d: r^2=0.006, p=.3).

Interim Discussion

By eliciting open-ended narratives, we found that having less power has a restrictive effect on the range of narratives produced. Low power narratives also tended to be self-rated by speakers as more informative and believable .

In the next experiments, we presented speakers with pre-written narrative choices. Narrative choices were written based on common responses from speakers, and were selected to include a range of narratives that were likely to vary in their informational, reputational, and presentational values (e.g., full truths vs. outright lies vs. evasions).

By asking speakers to rate the same set of narratives in terms of how likely they are to choose each narrative, as well as their informational, reputational, and presentational values, we can more directly test how power interacts with these goals to affect speakers' narrative choices. Previous results would suggest that low power speakers would place more importance on informational and presentational goals.

Further, by having all speakers rate the same set of narratives, we can try to separate the effect of power on how speakers *weigh* each goal versus evaluate their narratives. These two possibilities (effect of power on weights vs. values) are not mutually exclusive, but they cannot be teased apart with open-ended responses, since there was large variation in speakers' narratives.

Narrative Choice (Experiments 1b-d, 2a)

Participants

After writing their open-ended response, speakers from Experiments 1b-d were additionally shown pre-written narrative choices and asked to rate them. In an additional experiment (2a), speakers were only shown narrative choices (Low: n=145, Equal: n=145, High: n=151), and were recruited in the same way as described before.

Experimental Stimuli & Procedures

In Experiments 1b-d, narrative choices were always shown after speakers had written their own open-ended narratives. Within each experiment, speakers saw all the narrative choices , and were told that they were responses written by participants in a previous experiment. The number of narratives presented differed across experiments (in E1b: 6 narratives, E1c: 6 narratives, E1d: 3 narratives, E2a: 12 narratives), with some overlap across experiments with small variations in wording (all narratives are listed in **Figure 3**, lower right). For each narrative, speakers rated the likelihood of sharing the narrative with the listener ("How likely are you to say this to the [supervisor/co-worker/intern]?" Not at all likely - Very likely). They then rated each narrative on informational, reputational, and presentational values.

Results

Low Power speakers weight informational values more and reputational values less

Does power affect how speakers weigh informational, reputational, and presentational values when considering the utility of alternative narratives? To test this, we looked at the interacting effects of power and each value on choice ratings. This analysis collapsed data across experiments, to include as many narratives as possible (experiments and narratives were models as random effects, but the variance explained by experiments was low and taken out due to singular fit). As shown in **Fig. 3** (lower left, model estimates), high power speakers compared to low power speakers weighed informational values less and reputational values more, but no significant difference for weight on presentational value (interaction of power and info: B=-0.06, SE=0.02, p=.001; interaction of power of rep: B=0.07, SE=0.02, p=.004; interaction of power of pres: B=-0.00, SE=0.02, p=.99)

Further, the main effects of informational, reputational, and presentational values were significant (**Fig. 3** lower left) and explained more variance than models with each term taken out, validating the idea that speakers choose what to say by weighing their informational, reputational, and presentational goals.

Low Power speakers think uninformative and false narratives are likely to arouse more suspicion

Finally, we examine the effect of power on speakers' evaluations of narratives on informational, reputational, and presentational values (**Fig. 3**, top).

We found a significant effect of power on all three values (lower ratings overall for the low power group compared to the high power group, effect of power on informational value: B=1.91, SE=0.57, p<.001; reputational value: B=2.2, SE=0.59, p<.001; presentational value: B=2.2, SE=3.73, p<.001). Consistent with previous results, this held even for comparing low vs. equal power speakers.

We next explored how power affects ratings within each narrative (paired tests across low vs. high and low vs. equal groups). **Fig. 3** (top) shows ratings averaged across experiments and narratives. Grey highlights additionally

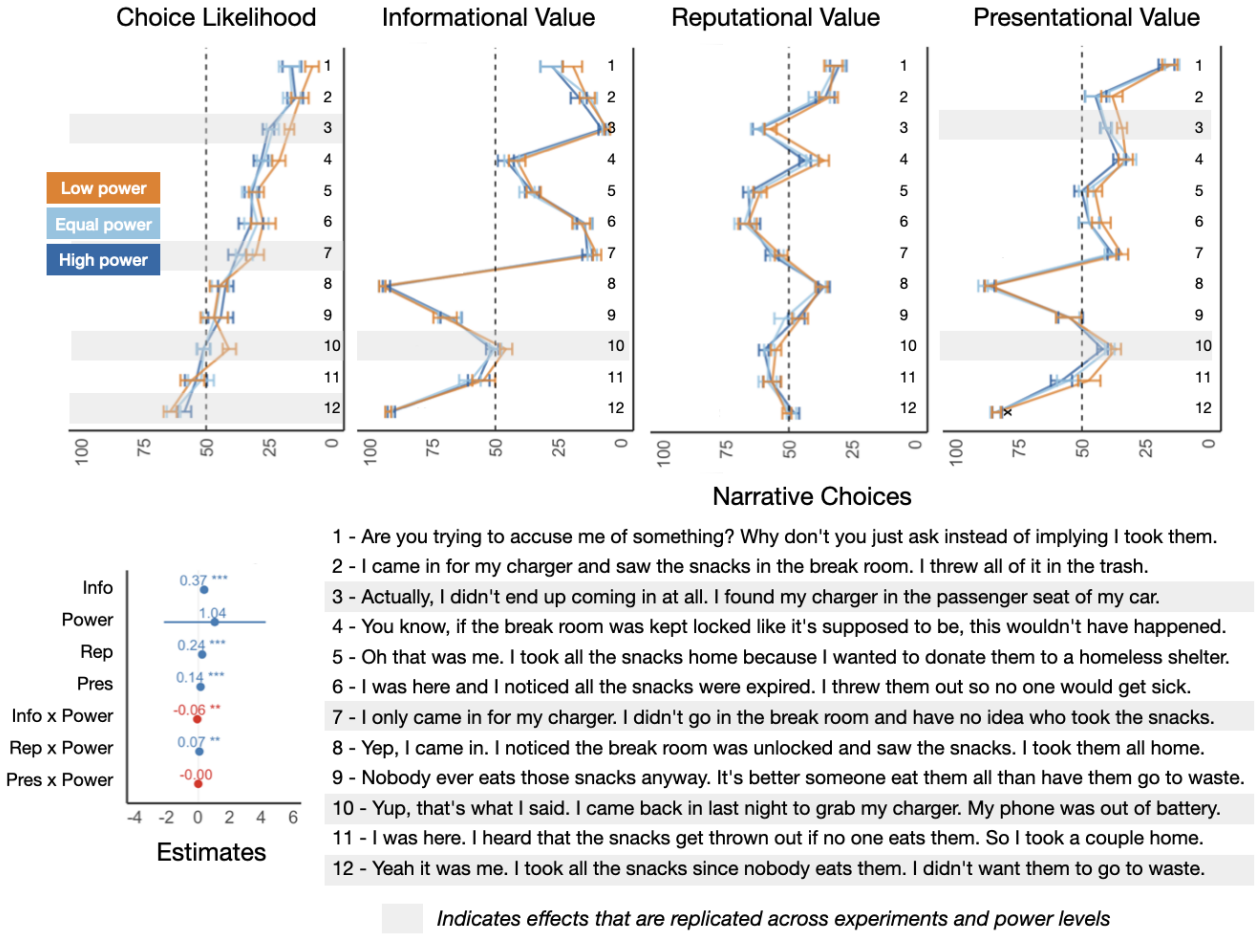

Figure 3. Top: Effect of power on informational, reputational, and presentational values of narrative choices (Experiments 1b-d, 2a). Bottom left: Model estimates for interacting effect of power and info/rep/pres values on choice ratings. Bottom right: narrative choices.

indicate narratives for which there is a consistent effect of power, findings that are replicated across experiments and across comparisons (both low vs. high and low vs. equal).

Consistent with the previous results, low power speakers show a stronger preference for the most complete and truthful narrative (#12). For all other narratives where there is an effect of power on evaluation of informational and presentational values, low power narrators are less likely to choose them (#10, 7, and 3).

Discussion

Across multiple studies, we find that low power leads to more similar narratives, demonstrating the idea that power restricts communicative freedom. Further, low power speakers cared more about the informational and presentational values of their narratives, and less about their reputational values. Exploration of particular narratives also confirmed that low power speakers preferred telling the truth, and were hesitant about narratives that are uninformative (either because they do not provide all the relevant information or are clearly false). Finally, we find evidence that power not only affects how speakers *weigh* their competing goals, it can also affect the valuation of narratives according to how well they meet each goal. Most notably, low power speakers thought there was a greater chance that their uninformative narratives would arouse suspicion.

The finding that power affects the space of choices available is of theoretical importance: current theories of language production assume that speakers decide what to say by comparing the utility of a small number of possible alternatives. What set of alternatives speakers actually consider, however, is typically not measured, and participants are typically given predetermined choices (Degen, 2023; but see also Degen & Tanenhaus, 2016; Gotzner & Romoli, 2022). The current results highlight that studies should be careful in assuming alternatives, especially for more complex utterances, and more work needs to be done to understand what kinds of factors can affect the space of alternatives speakers consider.

The generalizability of the finding that low power restricts narrative choices is an important open question. The reduction in choices follows from certain strategies (i.e., explicitly lying or being under-informative) becoming less favorable for low power speakers. More simply, low power speakers having fewer narrative options follows from the definition of power as power-to (act). Therefore, the same effect (restricted choices for low power) should in principle always hold where there is a power differential. Nonetheless, future work should try to replicate the effect in different contexts, and test whether low power speakers may sometimes become more likely to give varying utterances.

As others have pointed out, we expect *some* kinds of effects of power should vary greatly with context (Eelen, 2001; Locher, 2004; Harris, 2006; see also Qian et al., 2023). In fact, contrary to other studies, here we find that low power speakers become more informative, suggesting that power does not always lead people to choose a particular strategy (e.g., telling the truth). Rather, whether speakers place higher or lower importance on informational values might depend on the stakes and consequences of having less power, as well as the kind and source of power.

For instance, in our vignette, admitting a minor wrongdoing may not have been as costly as lying about it and being found out, leading many speakers to prefer to tell the truth. In other contexts, the costs could be reversed–e.g., if the wrongdoing is a crime, revealing the truth is extremely costly, and the effect of low power may lead speakers to lie even more. Furthermore, in many everyday situations, informational and reputational goals do not conflict (e.g., if telling the truth does not negatively affect the speaker's reputation). A recent paper finds that speakers not facing a goal conflict produce more homogenous narratives than those facing a conflict (Kim & Crockett, under review). We predict that power is much less likely to affect the range of choices speakers consider in low-conflict situations. In general, future work should systematically test the effect of power in a variety of contexts (e.g., beyond one vignette, the workplace, English).

Further work is needed to understand precisely why low power speakers assigned lower presentational values for some narratives, as there are multiple possible answers. Two narratives in particular were affected: an evasive and misleading narrative that does not provide relevant information (#10), and one that was clearly false and explicitly denied the speaker's wrongdoing (#3). The misleading was additionally rated low in informational value, which may reflect low power speakers anticipating their listener to have higher norm-based expectations about informativeness (e.g., a supervisor likely expects an employee to be complete and relevant in their answer, and become suspicious when the employee does not meet this norm). Lower presentational value for the false narrative may instead reflect low power speakers' increased possibility and fear of being found out. Another possibility is that low power speakers anticipated reduced credibility from listeners due to their lower position (similar to Carney et al., 2013; Ulatowska & Cislak, 2022).

Philosophers have argued that self-silencing (where speakers become less informative than they would otherwise) and extracted speech (where speakers become more informative) are both cases of *injustices*, as the speaker is being prevented from communicating in a way that they would be free to do if not for their lack of power (e.g., Kukla, 2014; Dotson, 2011; McKinney, 2016). There are many possible instances of discursive injustices in the real world. For instance, feeling the need to tell the truth could take the form of being pressured to disclose personal information (e.g., a disability). Likewise, feeling pressured to lie may range from lying about your opinion about a project, to making a false confession under duress, or giving "consent" to sex when you don't want to (Lackey, 2021; McKinney, 2016). Future work should study such cases in the real world that have particularly serious social and ethical implications.

Acknowledgments

The authors would like to thank Clara Colombatto and members of the Crockett Lab for helpful discussions of this research. This work was supported by the John Templeton Foundation grant on How Stories Inspire Virtuous Behavior (#61495 & #62816).

References

Banerjee, R., Heyman, G. D., & Lee, K. (2020). The development of moral self-presentation. *The Oxford handbook of moral development: An interdisciplinary perspective*, 92-109.

Bourdieu, P. (1977). Outline of a theory of practice. Cambridge University Press.

Brown, P., & Levinson, S. C. (1987). Politeness: Some universals in language usage. Cambridge University Press.

Camp, E. (2018). Insinuation, common ground, and the conversational record. *New work on speech acts*, *1*, 40-66.

Carney, D. R., Dubois, D., Nichiporuk, N., ten Brinke, L., Rucker, D. D., & Galinsky, A. D. (2013). The deception equilibrium: The powerful are better liars but the powerless are better lie-detectors. *Manuscript submitted for publication*.

Carmona, C. (2023). Silencing by not telling: testimonial void as a new kind of testimonial injustice. In *Testimonial Injustice and Trust* (pp. 35-50). Routledge.

Clementson, D. E. (2018). Deceptively dodging questions: A theoretical note on issues of perception and detection. *Discourse & Communication*, *12*(5), 478-496. 357–372. https://doi.org/10.1177/0261927X12468587

Dahl, R. A. (1957). The concept of power. *Behavioral science*, *2*(3), 201-215.

Davis, E. (2016). Typecasts, tokens, and spokespersons: A case for credibility excess as testimonial injustice. *Hypatia*, *31*(3), 485-501.

Degen, J. (2023). The rational speech act framework. *Annual Review of Linguistics*, *9*(1), 519-540.

Degen, J., & Tanenhaus, M. K. (2016). Availability of alternatives and the processing of scalar implicatures: A visual world eye-tracking study. *Cognitive science*, *40*(1), 172-201.

Deluga, R. J. (1995). The relation between trust in the supervisor and subordinate organizational citizenship behavior. *Military Psychology*, *7*(1), 1-16.

Dotson, K. (2011). Tracking epistemic violence, tracking practices of silencing. *Hypatia*, *26*(2), 236-257.

Eelen, G. (2001). A critique of politeness theories. St. Jerome Publishing.

Fiske, S. T. (2018). Controlling other people: The impact of power on stereotyping. In *Social cognition* (pp. 101-115). Routledge.

Foucault, M. (1971). Orders of discourse. *Social science information*, *10*(2), 7-30.

Frazier, M. L., & Bowler, W. M. (2015). Voice climate, supervisor undermining, and work outcomes: A group-level examination. *Journal of Management*, *41*(3), 841-863.

Fricker, M. (2007). Epistemic injustice: Power and the ethics of knowing. Oxford University Press.

Gotzner, N., & Romoli, J. (2022). Meaning and alternatives. *Annual Review of Linguistics*, *8*(1), 213-234.

Goodman, N. D., & Frank, M. C. (2016). Pragmatic language interpretation as probabilistic inference. *Trends in cognitive sciences*, *20*(11), 818-829.

Harris, S. (2003). Politeness and power: Making and responding to requests in institutional settings. *Text & Talk*, *23*(1), 27-52.

Harris, S. (2006). Politeness and power. In *The Routledge companion to sociolinguistics* (pp. 122-129). Routledge.

Hobbes, T. (1985). 1651. Leviathan. CB MacPherson, ed.

Holmes, J., & Stubbe, M. (2015). *Power and politeness in the workplace: A sociolinguistic analysis of talk at work*. Routledge.

Holtgraves, T. (1997). Yes, but... positive politeness in conversation arguments. *Journal of Language and Social Psychology*, *16*(2), 222-239. 278–292. https://doi.org/10.1177/0261927X970163003

Janssen, O., & Gao, L. (2015). Supervisory responsiveness and employee self-perceived status and voice behavior. *Journal of Management*, *41*(7), 1854-1872.

Keltner, D., Gruenfeld, D. H., & Anderson, C. (2003). Power, approach, and inhibition. *Psychological review*, *110*(2), 265.

Kim, J. S., Colombatto, C., & Crockett, M. J. (2024). Goal inference in moral narratives. *Cognition*, *251*, 105865.

Kim, J. S., & Crockett, M. J. (2025). Joint optimization of informational and reputational goals in moral narratives. *Manuscript under review.*

Kim, J., & Crockett, M. J. (2022). Narrating the" what" and" why" of our moral actions. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 44, No. 44).

Kukla, R. (2014). Performative force, convention, and discursive injustice. *Hypatia*, *29*(2), 440-457.

Lackey, J. (2021). False confessions and subverted agency. *Royal Institute of Philosophy Supplements*, *89*, 11-35.

Lakoff, R. (1989). The Limits of Politeness: Therapeutic and Courtroom Discourse in Linguistic Politeness II. *Multilingua*, *8*(2-3), 101-129.

Lindsey, L. L. M., Dunbar, N. E., & Russell, J. C. (2011). Risky business or managed event? Perceptions of power and deception in the workplace. *Journal of Organizational Culture, Communications and Conflict*, *15*(1), 55.

Locher, M. A. (2004). Power and politeness in action: Disagreements in oral communication. Mouton de Gruyter.

Mazzei, A., Ravazzani, S., Butera, A., & Fisichella, C. (2025). Do you dare to speak up at work? A study on organizational conditions enabling or discouraging employee voice. *Journal of Public Relations Research*, *37*(1-2), 115-134.

McKinney, R. A. (2016). Extracted speech. *Social Theory and Practice*, 258-284.

Medina, J. (2011). The relevance of credibility excess in a proportional view of epistemic injustice: Differential epistemic authority and the social imaginary. *Social*

Epistemology, *25*(1), 15-35. 15–35. https://doi.org/10.1080/02691728.2010.534568

Morand, D. A. (2000). Language and power: An empirical analysis of linguistic strategies used in superior–subordinate communication. *Journal of organizational behavior*, *21*(3), 235-248.

A. Morand, D. (2014). Using politeness to model the psychosocial dynamics of power in organizational interaction. *International Journal of Organizational Analysis*, *22*(2), 247-273.

Morrison, E. W. (2014). Employee voice and silence. *Annu. Rev. Organ. Psychol. Organ. Behav.*, *1*(1), 173-197.

Pfrombeck, J., Levin, C., Rucker, D. D., & Galinsky, A. D. (2022). The hierarchy of voice framework: The dynamic relationship between employee voice and social hierarchy. *Research in Organizational Behavior*, *42*, 100179.

Pitkin, H. F. (1972). Wittgenstein and justice: On the significance of Ludwig Wittgenstein for social and political thought. University of California Press.

Qian, P., Bridgers, S., Taliaferro, M., Parece, K., & Ullman, T. D. (2024). Ambivalence by design: A computational account of loopholes. *Cognition*, *252*, 105914.

Rogers, T., Zeckhauser, R., Gino, F., Norton, M. I., & Schweitzer, M. E. (2017). Artful paltering: The risks and rewards of using truthful statements to mislead others. *Journal of personality and social psychology*, *112*(3), 456.

Sedikides, C., Hoorens, V., & Dufner, M. (2015). Self-Enhancing Self-Presentation: Interpersonal, Relational, and Organizational Implications. In *Self-Concept, Motivation and Identity: Underpinning Success with Research and Practice* (pp. 29-55). Information Age Publishing.

Tangirala, S., & Ramanujam, R. (2008). Employee silence on critical work issues: The cross level effects of procedural justice climate. *Personnel psychology*, *61*(1), 37-68.

Thibaut, J. W., & Kelley, H. H. (1959). The social psychology of groups. Wiley.

Ulatowska, J., & Cislak, A. (2022). Power and lie detection. *PloS one* , *17* (6), e0269121.

Dyne, L. V., Ang, S., & Botero, I. C. (2003). Conceptualizing employee silence and employee voice as multidimensional constructs. *Journal of management studies*, *40*(6), 1359-1392.

Weber, M. (1978). *Economy and society: An outline of interpretive sociology* (Vol. 1). University of California press.

Weissman, B., & Terkourafi, M. (2019). Are false implicatures lies? An empirical investigation. *Mind & Language*, *34*(2), 221-246.

Yoon, E. J., Tessler, M. H., Goodman, N. D., & Frank, M. C. (2020). Polite speech emerges from competing social goals. *Open Mind*, *4*, 71-87.